

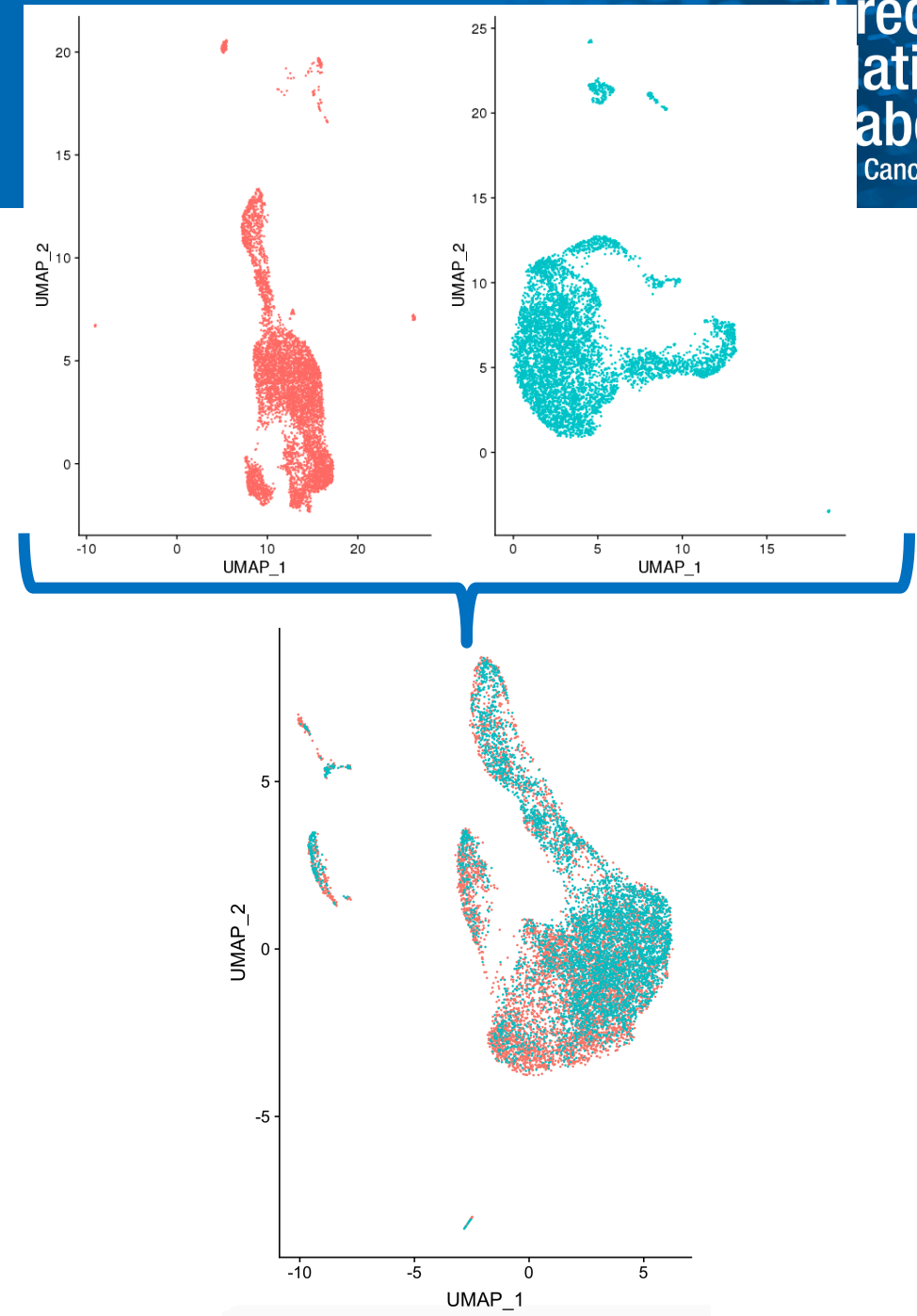
Multi-sample single cell analysis

Combining Samples

- **10X CellRanger can be used to combine samples for 10X captures**
 - Aggregate function
 - Generates UMI expression matrices, basic sample statistics, interactive analysis platform
 - Generates a summary report and a cloupe file that can be used for additional analysis
- **Other tools can combine samples downstream:**
 - Seurat
 - Scraper
 - Scanpy
 - Monocle
 - Etc.

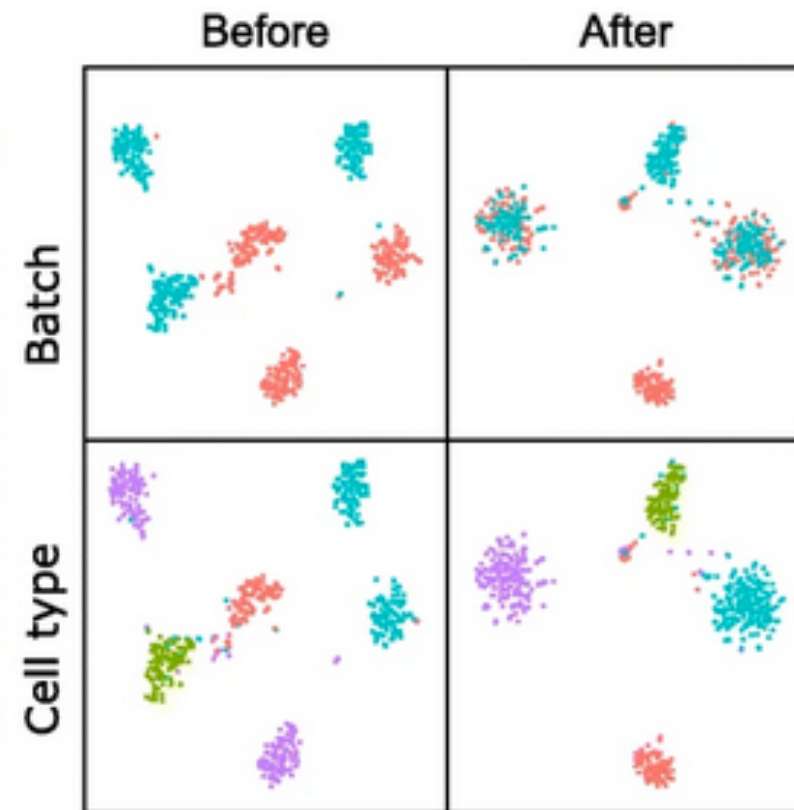
Combining samples

- Sample combination after sample QC is relatively straightforward
- It is strongly encouraged to rescale data following recommended practices
- This helps account for inter-sample variation when conducting downstream analysis



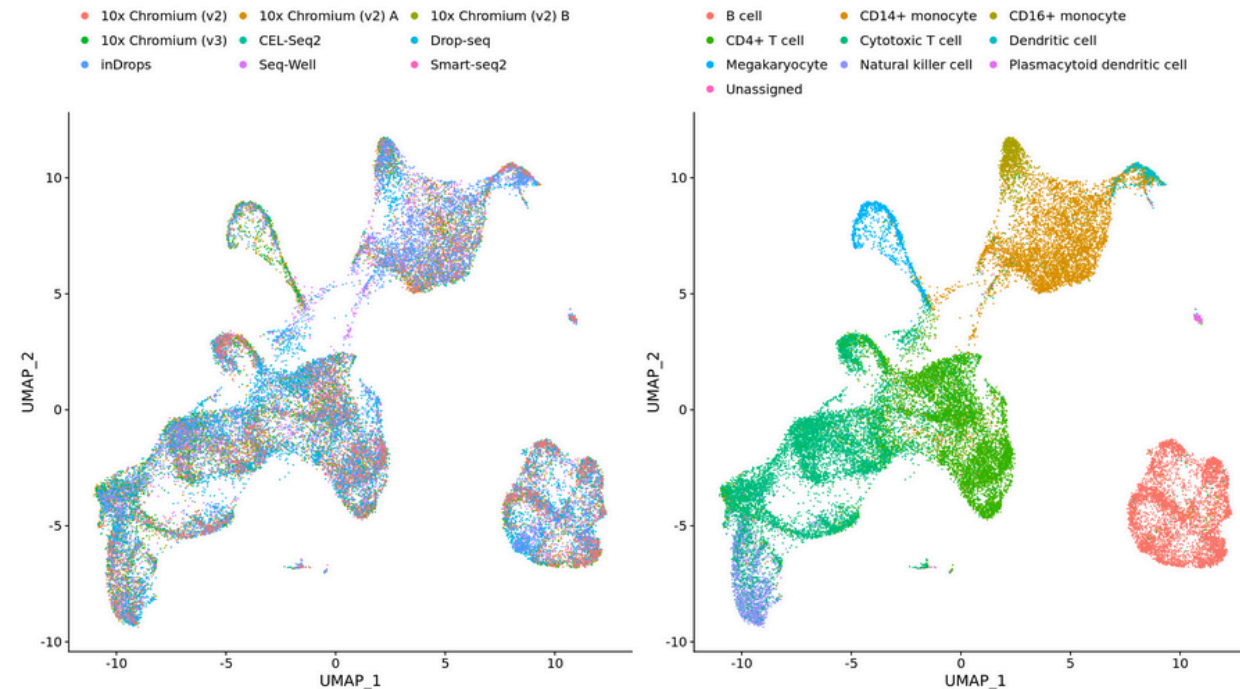
Batch correction: an overview

- When multiple samples and runs are combined, batch effects can occur
- This is reflective of any experiment that has multiple runs
- Multiple methods are available for batch correction
 - Seurat `integrate`
 - Scanorama
 - FastMNN
 - Harmony
 - Etc.



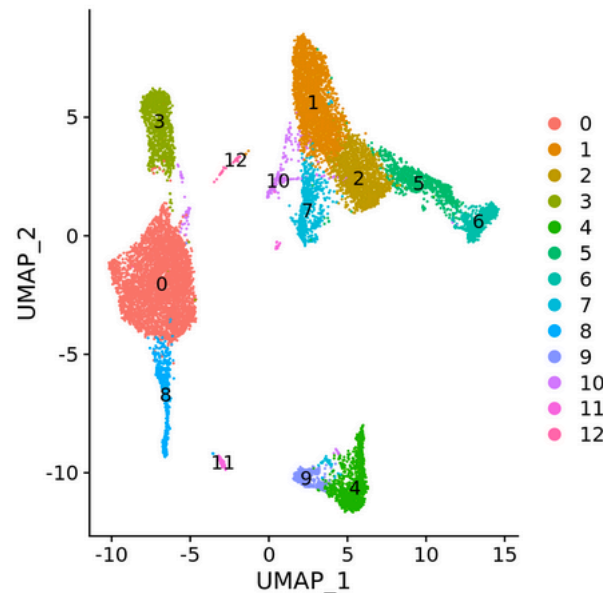
An example: Seurat and SCTransform

- Seurat v3 introduced SCTransform as a normalization technique that is tailored specifically to sample integration
- Much like finding features in single samples, integration identifies anchors in the most variable genes
- Ideally, the batch effects are minimized, allowing the biology to drive differences

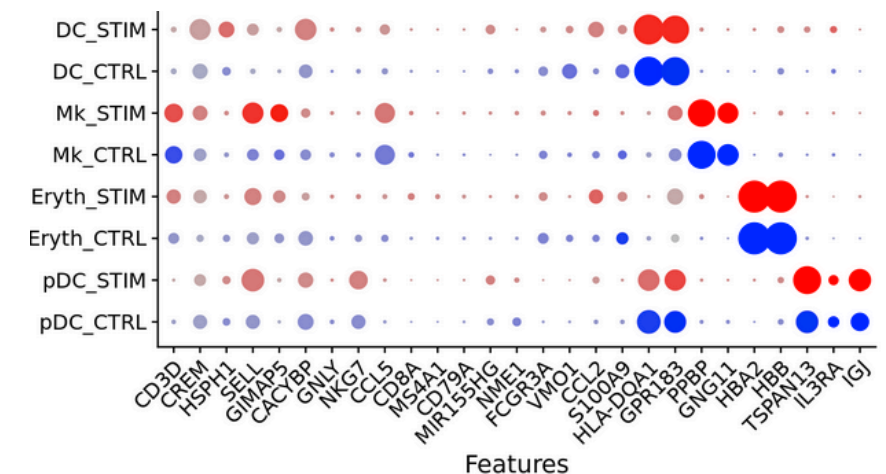
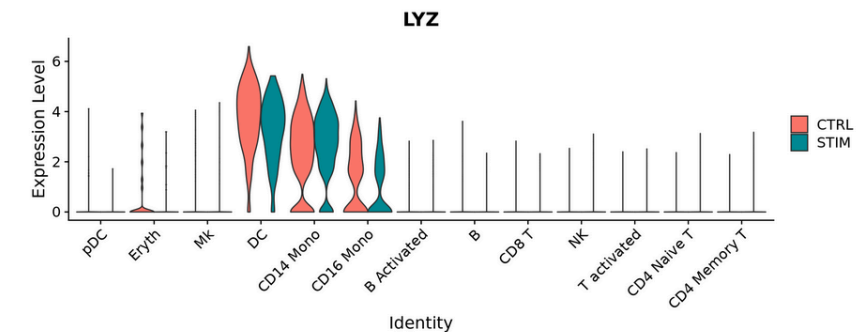


Multi-sample downstream analysis

- Many of the same analyses performed on individual samples can be conducted on multi-sample data
- This includes:
 - Clustering analysis
 - Differential expression
 - Pathway analysis
 - Trajectory analysis



##	CTRL_p_val	CTRL_avg_logFC	CTRL_pct.1	CTRL_pct.2	CTRL_p_val_adj
## GNLY	0	4.179377	0.944	0.045	0
## NKG7	0	3.163193	0.954	0.084	0
## GZMB	0	2.926436	0.841	0.043	0
## CLIC3	0	2.400545	0.599	0.024	0
## FGFBP2	0	2.242469	0.503	0.020	0
## CTSW	0	2.077683	0.533	0.029	0
##	STIM_p_val	STIM_avg_logFC	STIM_pct.1	STIM_pct.2	STIM_p_val_adj
## GNLY	0.000000e+00	4.077669	0.951	0.059	0.000000e+00
## NKG7	0.000000e+00	2.901784	0.957	0.080	0.000000e+00
## GZMB	0.000000e+00	3.124419	0.895	0.060	0.000000e+00
## CLIC3	0.000000e+00	2.446643	0.620	0.031	0.000000e+00
## FGFBP2	9.638295e-158	1.476625	0.256	0.016	1.35447e-153
## CTSW	0.000000e+00	2.180716	0.596	0.035	0.000000e+00
##	max_pval	minimump_p_val			
## GNLY	0.000000e+00	0			
## NKG7	0.000000e+00	0			
## GZMB	0.000000e+00	0			
## CLIC3	0.000000e+00	0			
## FGFBP2	9.638295e-158	0			
## CTSW	0.000000e+00	0			



Cell type annotation

- Identification of individual cell types can improve biological understanding of the data
- Cells can be annotated by hand with a list of known markers
- Tools can speed up the process by comparing cell transcription profiles to curated databases
 - SingleR
 - scPred
 - Garnett
 - scMatch

